The Logic of Comparative Theory Evaluation

by M.L.G. Redhead, Chelsea College, London.

25.10.75

Abstract

Schaffner's logic of comparative theory evaluation is criticised for an incorrect interpretation of Bayes's theorem. A corrected analysis is given and extended to the case of ~~a number~~ multiple successful predictions by a theory. The application of the method to the appraisal of quantitative predictions is discussed with examples. ~~Finally two hypothetical worlds are presented for the special case in which successful and unsuccessful predictions are interspersed in the development of a research programme and a simple representation for the life and death of a research programme is provided.~~

A Bayesian model is ~~the~~ next presented for the general case in which successful and unsuccessful predictions are interspersed in the development of a research programme. An alternative approach incorporating the notion of methodological decision is discussed and simple representations are provided for the life and death of a research programme.

The Logic of Comparative Theory Evaluation

1. A Critique of Schaffner's Account

In a recent note Schaffner ([1974]) has ~~[discussed]~~ given a formal discussion of the notion of adhocness in terms of a Bayesian model for the appraisal of theories. Schaffner develops his general ideas in the context of a critique of Zahar's [1973] which was concerned with the particular problem of comparing the Einstein and Lorentz research programmes. Zahar suggests the following analyses of adhocness:*

[illegible] ad hoc$_1$: A theory is said to be ad hoc$_1$ if it has no novel consequences as compared with its predecessor.

ad hoc$_2$: [A theory] ... is ad hoc$_2$ if none of its novel predictions have been actually 'verified'.

ad hoc$_3$: [A] -- theory is said to be ad hoc$_3$ if it is obtained from its predecessor ~~[illegible]~~ through a modification of ~~[illegible]~~ the auxiliary hypotheses which does not accord with the spirit of the heuristic of the programme.

[illegible] Zahar explains the meaning of novelty as follows:#

A fact will be considered novel with respect to a given hypothesis ~~[illegible]~~ if it did not belong to the problem-situation which generated the construction of the hypothesis.

[illegible] Schaffner's attempted elucidation is really rather confused. He first discusses the notions of ad hoc$_1$ and ad hoc$_3$. The

---

*[1] Zahar [1973] p. 101 #[2] Zahar [1973] p. 103

first he describes as a logical dream, since the novel consequences of a theory cannot in practice be "surveyed", so the question of whether a theory is ad hoc, can only be discussed relative to the extent to which novel consequences have been looked for at the particular epoch of the evaluation. Ad hoc, Schaffner claims to be "vague to the point of inapplicability". For Schaffner ad hoc is "close to the sense in which ad hoc is used in science" and he announces that his Bayesian analysis will be brought to bear on this sense of ad hoc.

He begins his analysis[3] by referring to the general division theorem by which he actually means Bayes's theorem. Schaffner says that this is a noncontroversial theorem in the probability calculus and refers to Salmon's [1967] account.

He claims in a footnote[4] to have improved on Salmon's presentation. Unfortunately he so far "improves" the theorem that he actually renders it false.

new para Denoting by $p(T/b\&e)$ the probability of a theory T to be true in the light of background knowledge b and the positive outcome e of some experiment as part of b, by $p(T/b)$ the prior assessment of T, by $p(e/T\&b)$ the probability of obtaining e given T and b, and by $p(e/b)$ "the probability of obtaining e on the basis of theories other than T", Schaffner states

[3] Schaffner [1976] p. 69–73.

[4] Ibid Footnote 3 p. 70.

(1) $$p(T|h\,\&\,e) = \frac{p(T|h)\cdot p(e|T\,\&\,h)}{p(e|h)}$$ (A)

In a footnote[5] he explains $p(e|h)$ can be expanded in terms of a set of n "alternative competing theories" (my italics) in the form

(2) $$p(e|h) = \sum_{i=1}^{n} p(T_i|h)\, p(e|T_i\,\&\,h).$$

now pard

Now (1) is of course true only if $p(e|h)$ includes a contribution from T itself, which Schaffer explicitly denies ~~This confus~~ in his defenition of $p(e|h)$. This unfortunate oversight now makes nonsense of Schaffer's subsequent analysis. Because he assumes $p(e|h)$ is independent of $p(T|h)$ he makes his criterion for a theory to be ad hoc[6] (footnote) Schaffer actually works in terms of hypotheses being ad hoc but we want to keep the argument as simple as possible) depend essentially on the smallness of $p(T|h)$. The argument[7] for $p(T|h)$ being small is essentially that it has no "theoretical support". This looks suspiciously like a reference to ad hoc$_3$. and Schaffer now appears to be claiming[8], that a theory is ad hoc$_2$ in virtue of its being ad hoc$_3$. So he is not really giving an independent analysis of ad hoc$_2$ at all. But it is further to critiqe his detailed arguments given they are all derived by using

5 Quid Footnote 2 p. 70.

Insertion
new Ford

Note that we follow Schaffner in interpreting $p(A/B)$ as probability that A is true given that B is true.

Note also that e is supposed in quotation (3) to refer to a prediction made by the theory T (we could perhaps write e_T to emphasize this important point) so $p(e/T\&C)$ means the probability that the prediction e_T derived from the theory T is a true prediction given that the background information C is true but the theory T is false (i.e. its consequences are not guaranteed to be true although "by accident" they may be true)

the correct formulation of Bayes's theorem. The only statement which Scheffler gets right is where he discusses the "odds form" for comparing the evaluation of a theory with its rivals*7. Here in taking the appropriate ratio $p(e/h)$ cancels out thus is the one point at which his argument is not invalidated by his incorrect analysis of $p(e/h)$.

2. The Bayesian Analysis of Adhocness

It is in fact a fairly simple matter to get an "internal" analysis of adhoc by using the correct form of Bayes's theorem, which we write as

$$(3)\quad P(T/h\&e) = \frac{P(T/h)\cdot P(e/T\&h)}{P(e/T\&h)\,P(T/h) + P(e/\sim T\&h)\,P(\sim T/h)}$$

where $\sim T$ denotes the negation of T.

Inequations

Writing $P(T/h) = x$, $P(e/\sim T\&h) = \varepsilon$

and taking $P(e/T\&h) = 1$ and using $P(\sim T/h) = 1 - x$ yields

$$(4)\quad P(T/h\&e) = \frac{x}{x + \varepsilon(1-x)}$$

We define an enhancement rate Ω by

$$\Omega = \frac{P(T/h\&e)}{P(T/h)}$$

where using (4)

*7. Ibid p 72.

we obtain the simple result

$$(5) \qquad \Omega = \frac{1}{x + \varepsilon(1-x)} .$$

new para

We can now explain that a theory T is ad hoc with respect to the experiment e if $\varepsilon = 1$, i.e. if the explanation of e by T in no way depends on the truth or falsehood of T, both of which eventualities lead ~~to~~ with certainty to the result e. This is just what a scientist means when he says T was an ad hoc explanation of e, namely T was devised for the express purpose of explaining e, so the explanation of e is guaranteed independent of whether T is true or false.

To show the consistency of our analysis if we put $\varepsilon = 1$ in (5) we get $\Omega = 1$ so the posterior and prior probabilities of T are equal (there is no enhancement) and this again is just what we expect from an ad hoc explanation of e, namely e ~~gives us no~~ itself gives us no additional information for assessing the truth of T.

new para

We notice that if $x \leq \varepsilon \leq 1$, then $\Omega \simeq 1/\varepsilon$ so in this case we get a big enhancement ~~of~~ and the theory is

for from being ad hoc, this is just the situation in which Schaffer would claim that the theory was ad hoc, which highlights the way in which his analysis goes astray.

3. ~~The case of~~ Multiple Predictions

To develop our analysis a little further we can consider how a theory builds up a favourable approval as it makes a number of successful predictions $e_1, e_2 \ldots e_n$ say. Denote by P_s the posterior probability $P(T/\&e_1 \& e_2 \ldots \& e_s)$ after s successful predictions. Assuming the predictions are quite independent and ~~[associated with e]~~ for simplicity are all associated with the same value of ε, we can clearly write

$$P_n = \Omega^{(n)} \times \Omega^{(n-1)} \times \ldots \Omega^{(1)} \cdot P_0$$

where $P_0 = x$ according to our previous notation

and
$$\left.\begin{aligned} \Omega^{(s)} &= \frac{1}{P_{s-1}(1-\varepsilon)+\varepsilon} \\ P_{s+1} &= \Omega^{(s+1)} \cdot P_s \end{aligned}\right\}$$

new hard The solution of this recursion is by inspection (or more simply by replacing $\varepsilon \to \varepsilon^n$ in the formula for P_1 (see (4) above). We obtain

$$(6) \qquad P_n = \frac{1}{1-\varepsilon^n + \varepsilon^n/x}$$

new hard We can also ask what is the probability for ~~obtaining~~ the $(n+1)^{th}$ prediction e_{n+1} being correct if the theory has already made n

successful predictions $e_1 \ldots e_n$. Denoting this probability by $P(e_{n+1})$ we easily obtain the result

(6) $$P(e_{n+1}) = \varepsilon + \frac{1-\varepsilon}{1-\varepsilon^n + \varepsilon^n/x}$$

If ε is a small quantity (i.e. $\ll 1$) ~~this~~ which will be the case if T is non-ad hoc with respect to all the predictions, we can write the following approximate formulae which will be perfectly satisfactory for the subsequent discussion

(7) $$P_n \simeq \frac{1}{1+\varepsilon^n/x}$$

(8) $$P(e_{n+1}) \simeq \varepsilon + P_n$$

(9) $$\Omega^{(n)} \simeq \frac{1}{\varepsilon + P_{n-1}}$$

We note the following features

(1) The value of P_n depends critically on the ratio ε^n/x.
For $\varepsilon^n/x \gg 1$, $P_n \ll 1$
and for $\varepsilon^n/x \ll 1$, $P_n \simeq 1$

~~At the~~ critical point $\varepsilon^n = x$, and there $P_n \simeq 1/2$.
So if initially $x \ll \varepsilon$ as n increases P_n will rise steeply as we reach the value $n = \ln x/\ln \varepsilon$.

(2) So long as $P_n \lesssim \varepsilon$, we have $P(e_{n+1}) \simeq \varepsilon$, but as P_n builds up towards unity, so does $P(e_{n+1})$

(3) So long as $P_{n-1} \lesssim \varepsilon$, $\lambda^{(n)} \simeq 1/\varepsilon$, but as P_{n-1} builds up towards unity the enhancement factor $\lambda^{(n)}$ also tends to unity

new para

To take a concrete ~~illustration~~ case we illustrate in Fig 1. P_n and $P(e_{n+1})$ as functions of n for the particular choice $x = .01$, $\varepsilon = .1$.

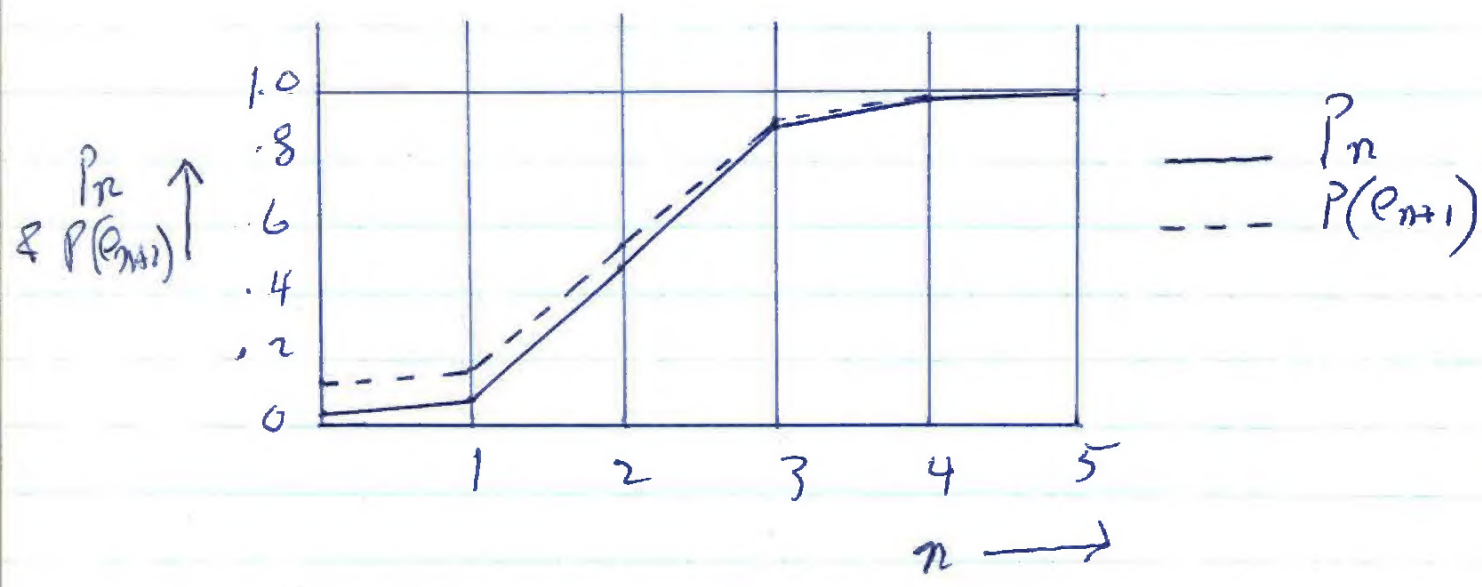


Fig 1.

4. Quantitative Predictions

We can apply this analysis to the important case of quantitative predictions.

Suppose a theory T predicts exactly an experimental result which is known to an accuracy of n significant figures. We assume as part of our background

knowledge is that the order of magnitude of the result is known, i.e. we disregard the prediction of zeros occurring before or after the significant figures in the experimental result.

As a concrete illustration we cite the theoretical predictions of quantum electrodynamics of the anomaly in the hydrogen spectrum known as the Lamb shift and the anomaly in the magnetic moment of the electron. For example the latter is now known experimentally to be 0.00115965 7.7 ± 3.5 Bohr magnetons whereas the theoretical prediction is 0.00115965 5.4 ± 3.3 magnetons. (For a good account of the flourishing agreement between theory and experiment for this phenomenon the reviews by Lautrup, Petermann and Sommerfield in their [1972] and de Rafael in their [1972] may be consulted.)

Clearly if we regard the prediction of each significant figure as an independent event then the appropriate value to take for ε is .1 since a false theory would have ten equal possibilities for failing in each digit. The question of what value to take for x is somewhat arbitrary. In agreement with Schaffner we do not follow a purely logical approach and set $x = 0$. For our purposes x reflects the scientists belief confidence in the new theory T. One could argue that a scientist would not spend great efforts developing the consequences of a theory he did not believe in. By analogy with the situation

in the Bayesian analysis of significance testing (see for example Pollard [1974]) we could take $x = 1/2$. Perhaps more realistically we should take x around 0.01 and adopt the sociological rule that when a scientist has a one percent level of confidence in the truth of his theory he would not seriously investigate it.

With this choice of parameters we see that the build up of confidence in a theory which makes correct quantitative predictions, as the accuracy of the experiment increases, would be illustrated by the graph for P_n in Fig. 1.

Of course $P(E_{n+1})$ is only given by this analysis so long as other factors which could potentially influence the result are known not to be significant. For a certain value of n this condition will fail. For example in the case of the anomalous magnetic moment of the electron the effect of hadronic couplings introduces theoretical uncertainties which lower ultimately make the prediction of the theory unreliable.*8

On the other hand the Lamb shift has so far been measured with an accuracy less by a factor of twenty so the physicist's confidence in refined quantitative predictions here is very high.

---

*8 According to Rich, Werley [1972] the hadron hadronic contributions to the electron anomaly would affect the tenth significant figure.

5. A Bayesian Model for the Methodology of Research Programmes

We now want to extend our analysis to the case where the theory T consists of two parts $T = c \,\&\, B$ where c is the hard core and B an auxiliary hypothesis selected from a set $\{B_i\}$ which represents the possible articulations of the protective belt in Lakatos's terminology (see for example Rus [1970])

We write

$$P(c/e) = \frac{P(e/c)\,P(c)}{P(e/c)\,P(c) + P(e/\sim c)\,P(\sim c)}$$

and expand $P(e/c) = P(e/c \,\&\, B)\,P(B/c)$ [9]

$$+\, P(e/c \,\&\, \sim B)\,P(\sim B/c)$$

(Footnote 9. This method of analysing Bayes's theorem due to ~~as suggested by~~ Jon Dorling)

Introducing abbreviations $P(c) = x$, $P(B/c) = y$,

$$P(e/c \,\&\, \sim B) = k_1,$$

$$P(e/\sim c) = k_2$$

and setting $P(e/c \,\&\, B) = 1$, $P(\sim B/c) = 1-y$, $P(\sim c) = 1-x$

then enhancement ratio for hard core is Ω_c

where $\Omega_c = \dfrac{P(c/e)}{x} = \dfrac{y + k_1(1-y)}{(y + k_1(1-y))x + k_2(1-x)}$

If we take $k_1 = k_2 = k$ say we get

(10) $$\Omega_c = \frac{y + k(1-y)}{(y + k(1-y))x + k(1-x)} = \frac{y + k(1-y)}{xy + k(1-xy)}$$

the enhancement ratio for the auxiliary hypothesis

Similarly for auxiliaries Ω_B, is of the form

(11) $$\Omega_B = \frac{x + k(1-x)}{(x + k(1-x))y + k(1-y)} = \frac{x + k(1-x)}{xy + k(1-xy)}$$

If we assume $P(e/\sim C \& B) = k$
and $P(e/\sim B) = k$.

for simplicity.

For an ad hoc prediction we take $k = 1$

to get $\Lambda_C = 1 = \Lambda_B$

so again there is no enhancement in this case.

For a non-ad hoc prediction take $k \leq xy$.

then $\Lambda_C \leq \dfrac{y}{yx} \leq 1/x$.

Similarly $\Lambda_B \leq 1/y$

so we now get big enhancements leading in our approximation limit to $P(e/e) = 1$ and $P(B/e) = 1$ after confirmation.

this is true <u>irrespective of magnitude of x and y.</u>

rewrite

We turn now to the interesting case of refutation and write

$$P(C/\sim e) = \frac{P(\sim e/C)P(C)}{P(\sim e/C)P(C) + P(\sim e/\sim C)P(\sim C)}$$

where $P(\sim e/C) = P(\sim e/C \& B)P(B/C) + P(\sim e/C \& \sim B)P(\sim B/C)$

and now substitute

$$P(re|c\&B) = 0$$

$$P(re|_2e) = 1-k_2, \quad P(re|c\&_2\beta) = 1-k_1$$

to give us then each, again assuming $k_1 = k_2 = k$

$$(12)\ \Omega_c = \frac{(1-k)(1-y)}{(1-k)(1-y)x + (1-k)(1-x)} = \frac{1-y}{(1-y)x+(1-x)} = \frac{1-y}{1-xy}$$

and

$$(13)\ \Omega_B = \frac{(1-k)(1-x)}{(1-k)(1-x)y + (1-k)(1-y)} = \frac{1-x}{(1-x)y+(1-y)} = \frac{1-x}{1-xy}.$$

~~[Take just the actual case that $k=1$~~

~~Then $\Omega_c =$~~] Thus formulae for Ω_c and Ω_B are independent of the value of k.

~~Consider $\Omega_c < 1$ if~~

~~$1-y < (1-y)x + (1-x)$~~

~~or $(1-y)(1-x) < 1-x$ which is always true since $1-y < 1$~~

~~Hence we have for reputation $\Omega_c < 1$~~

We note at once that for reputation $\Omega_c > \Omega_B$ are always less than unity, as we would expect.

* Footnote

10. The asymmetric effect of rotation in the Bayesian model was pointed out to me by Jon Dorling who was responsible for the ideas behind the preceding analysis

of successful and unsuccessful predictions.

In Fig. 2 ~~we illustrate~~ we illustrate the following sequence — two alternating successes and failures followed by a run of failures when we have assumed as parameter values $x = 0.9$, $y = 0.5$ and $h = 0.1$. The probability P_C for the hard core and the probability for the current articulation of the protective belt P_B are both shown. After each refutation we take $P_B = 0.5$ again for the new ~~optional~~ auxiliary hypotheses introduced to account for the anomaly.

We are assuming here the simple situation in which each articulation of the protective belt is triggered by an anomaly. Lakatos of course stresses in his [1972] that robustic research programmes can progress in an ocean of anomalies which are ignored, the articulation of the protective belt being carried along by the internal momentum of the positive heuristic of the programme. Our 'rational reconstruction' of Lakatos's methodology, like all rational reconstructions, is a ~~characterisation~~ caricature ~~(compare Lakatos [1972])~~

But suppose we have x near to 1 and y small then ~~$\Omega_e = \frac{1}{x}$~~ for $x=1$, $y=0$ we get

$$\Omega_e = 1$$

thus so no law of conflicts is chosen here.

But ~~if x is~~ is same here

$\Omega_B = 0$, so Refutation acts ~~symmetrically~~ asymmetrically on C and B provided we take x large, y small as reflecting the methodology of research programmes.[10] (Formal ~~to~~ * insertion)

new para

Note in particular we can complicate refutation without having to assume ξ & h small.

new para

The life and death of a research programme can now be modelled using our two formulae (10) & (12) to allow for the effects

For a successful prediction

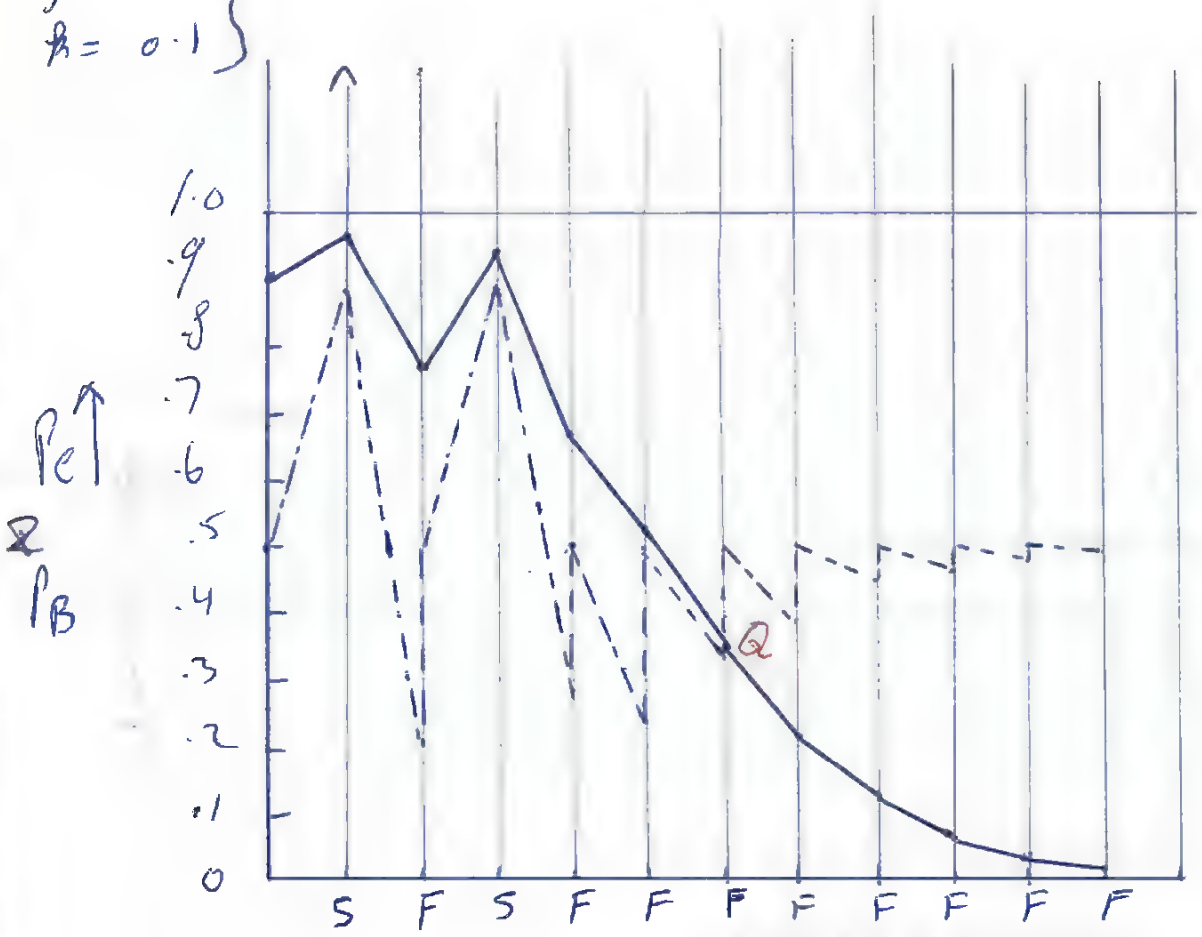
(14) $$\Omega_e = \cancel{\frac{y+h(1-y)}{(y+h(1-y))x+h(1-x)}} = \frac{y+h(1-y)}{xy+h(1-xy)}$$

and for a refutation

(15) $$\Omega_e = \cancel{\frac{1-y}{(1-y)x+1-x}} \quad \frac{1-y}{1-xy}$$

P.T.O.

$x = 0.9$
$y = 0.5$
$h = 0.1$



— P_C
--- P_B

Fig. 2.

S = successful production
F = failed production

new para
Suppose after fourth failure in the degeneration sequence P_C has fallen to 0.21 and that a success.

P_C must then recover to the value 0.60.

Even such P_C drop to .06 at the fourth failure in the degeneration sequence, a success would lift P_C to 0.26.

new para
At the "crossover" point of the P_C and P_B curves (marked Q in Fig 2) it would be more rational to change the hard core rather than the protective belt, i.e. we see here an internal rational criterion for abandoning a research programme, independent of any competing programme.

new page

At first sight we might suppose that if we have a run of successful predictions followed by a failure then the asymmetry in the effect of failure on P_c and P_B which was large initially would be eliminated as P_c and P_B were both now increased to values very close to unity. We shall now show that this rather natural assumption is unwarranted and that in our model the asymmetric effect of failure on P_c and P_B is preserved however many successful predictions precede the failure.

What we shall prove is the following:

After n successful predictions let ~~each~~ P_c increase from $P_c(0)$ to $P_c(n)$ and P_B from $P_B(0)$ to $P_B(n)$

Let the effect of a refutation on P_B, P_c after n successful predictions be ~~denoted by~~ denoted by $\Omega_c(n)$, $\Omega_B(n)$

Then we prove $\dfrac{\Omega_B(n)}{\Omega_c(n)} = \dfrac{\Omega_B(0)}{\Omega_c(0)} = \dfrac{1-P_c(0)}{1-P_B(0)}$

a result which is independent of n

~~To prove this result~~

We have at once the result

$$\frac{\Omega_B(n)}{\Omega_c(n)} = \frac{1-P_c(n)}{1-P_B(n)} \qquad \left(\text{see eqn } (15)\right)$$

Now use eqn (10) to obtain the result

$$P_c(n+1) = \frac{P_B(n) + P_B\left(1-P_B(n)\right)}{P_c(n)P_B(n) + P_B\left(1-P_c(n)P_B(n)\right)} \times P_c(n)$$

$$\cancel{P_e(n+1) = \frac{P_B(n)P_C(n) + P_e(n) - P_B(n)P_C(n)}{P_B(n)P_C(n) + 1 - P_B(n)P_C(n)}}$$

So $$1 - P_e(n+1) = \frac{P_e(1 - P_e(n))}{P_B(n)P_C(n) + P_e(1 - P_B(n)P_C(n))}$$

Similarly we can show

$$1 - P_B(n+1) = \frac{P_e(1 - P_B(n))}{P_B(n)P_C(n) + P_e(1 - P_B(n)P_C(n))}$$

$$\frac{1 - P_e(n+1)}{1 - P_B(n+1)} = \frac{1 - P_e(n)}{1 - P_B(n)}$$

which applied recursively gives the result

$$\frac{1 - P_e(n)}{1 - P_B(n)} = \frac{1 - P_e(0)}{1 - P_B(0)} = \frac{1 - x}{1 - y}$$

in our old notation

which is independent of n.

Hence we get for rotation after n seconds

$$\frac{\Omega_B(n)}{\Omega_e(n)} = \frac{1 - x}{1 - y} \quad \left(= \frac{1 - P_e(0)}{1 - P_B(0)} \right)$$

For our example $x = .9$, $y = .5$

so $\frac{\Omega_B(n)}{\Omega_e(n)} = 1/5$ and

effect of rotation on B is five times as high

on on C for all values of n,

although of course it is also true that $P_B(n)$ and $P_C(n)$ both tend rapidly towards 1 with increasing n.

new para

We stress that in this model the cognitive effect of refutation on probabilities let B, hard core C is directly accounted for at all stages by the expected value of P_C being much higher (closer to unity) than we assume for P_B.

6 The Methodological Decision ~~as an Alternative Approach~~

In ~~this~~ the model we have just described the effect of refutation on P_C is given a rational (Bayesian) basis.

But other approaches to the methodology of theory evaluation can best be understood as a ~~rather unhappy~~ mixture of rational and methodological decisions. We shall now illustrate some of these ideas.

new para

Thus faced with an anomaly there are various methodological prescriptions which tell the scientist how to proceed.

~~So far we have talked about successful predictions, but what happens when a theory makes an unsuccessful prediction, or what we have called an anomaly? There are various methodologies which tell the theorist how to deal with anomalies. In summary there is a crude Popperian picture which tells him to reject a theory if it is refuted. But the tenacity of theories is reflected in the Lakatos' Methodology of Research Programmes.~~

~~To analyze this situation, on the context of our Bayesian model again, we have~~ again wrote T as a conjunction of a hard core C and an auxiliary hypothesis B_1, which belongs to a set $\{B_i\}$ of adjustable hypotheses in the protective belt, adjustable that is in accordance with the positive heuristic of the research programme.

Thus $T = C \& B_1$.

Now suppose the theory T makes a false prediction e_1. Since this is impossible if the theory T were true we must write $P(e_1 | T \& u) = 0$ in equation (3) which leads immediately to $P(T | u \& e_1) = 0$. If we assume C and B_1 are independent hypotheses for the sake of simplicity, then we can write

$$P(T | u \& e_1) = P(C | u \& e_1) \times P(B_1 | u \& e_1)$$

So we are led to the condition

$$P(C | u \& e_1) = 0 \quad \text{or} \quad P(B_1 | u \& e_1) = 0$$

Which alternative we take cannot be decided by the probability calculus alone. It is a matter for methodological with the information: assumed so far alone at our disposal.

decision. Thus the Lakatos methodology can be expressed very simply. We choose to take $P(B_1 | h \& e_1) = 0$ and leave $P(c | h \& e_1)$ unchanged in the face of a false prediction e_1. We now invoke a new hypothesis B_2 which enables $c \& B_2$ to explain e_1. This is an ad hoc manoeuvre for $c \& B_2$ with respect to e_1 so $P(c \& B_2 | h \& e_1)$ can only be enhanced by new successful novel predictions. Suppose e_2 is successfully predicted by $c \& B_2$. Then our previous analysis gives

(14) $$P(c \& B_2 | h \& e_1 \& e_2)$$
$$= \Omega \times P(c \& B_2 | h \& e_1)$$

where Ω is a suitable enhancement factor.

We can rewrite (14) as:

$$P(c | h \& e_1 \& e_2) \times P(B_2 | h \& e_1 \& e_2)$$

(15) $$= \Omega \times P(c | h \& e_1) \times P(B_2 | h \& e_1)$$

But what we are actually interested in is how the credibility of the hard core behaves as the research programme progresses. To answer this question we must "distribute" the enhancement ratio between the hard core and the protective belt. But again the probability calculus alone cannot

answer this question. Again we have to make a methodological decision. But this decision must be constrained so as not to transgress the rules of the probability calculus. The situation is rather like Popper's formal account of concepts like verisimilitude. We seek a simple formula that satisfies certain conditions. We formulate these conditions for an inducum as follows.

next pages In a suitably abbreviated notation we let A & B stand for two probabilities and we write

$$\Omega\, AB = (\Omega_A \times A) + (\Omega_B \times B)$$

So the enhancement factor Ω ~~for A & B~~ is divided between A & B according to the factors Ω_A and Ω_B which will be functions of Ω, as well as both depending on A & B.

Regarding Ω_A & Ω_B as functions of Ω we list the following ~~natural~~ restrictions, assuming $A \geq B$ without loss of generality

(1) $\Omega_A(1) = \Omega_B(1) = 1$

(2) $\Omega_A\left(\frac{1}{AB}\right) = 1/A$, $\Omega_B\left(\frac{1}{AB}\right) = 1/B$

(3) $\Omega_A(\Omega)$ and $\Omega_B(\Omega)$ are monotonically increasing functions of Ω on the range $(1, 1/AB)$

(4.) $\Omega_A \times A > \Omega_B \times B$ ~~if A > B~~ for all Ω in range $(1, 1/AB)$

(5.) $\Omega_A + \Omega_B = \Omega$

(6) Writing $\Omega_A(\Omega) = f(A, B, \Omega)$

$$f(A, B, \Omega'\Omega'') = f(A, B, \Omega') \times f\left[f(A, B, \Omega')A, \frac{\Omega}{f(A, B, \Omega')}, B, \Omega''\right]$$

which similar condition on $\Omega_B(\Omega)$

Condition (2) arises from noting that the maximum value of Ω is $1/AB$ since we require that ΩAB is a probability, the maximum value for Ω_A is $1/A$ since $\Omega_A A$ is to be a probability, and similarly for Ω_B.

Condition (4) means that enforcement never leads to a "over-ever" in the assessment of A and B.

Condition (6) means that our rule for arriving at Ω_A can be built up consistently in stages.

Conditions (1), (3) and (5) require no comment.

It is perhaps surprising that a simple formula for Ω_A can be obtained satisfying all these conditions. viz

(18) $$\Omega_A = \cancel{\frac{\Omega(1-AB)}{1-A+B(1-A)\Omega}} \quad \frac{\Omega(1-AB)}{1-A+A(1-B)\Omega}.$$

Ω_B is then given trivially as Ω/Ω_A.

(17) i.e. $$\Omega_B = \cancel{\frac{1-B+B(1-A)A}{1-AB}} \quad \frac{1-A+A(1-B)\Omega}{1-AB}$$

For $A < B$ we interchange the formulae for Ω_A, Ω_B, at the same time exchanging A and B in these formulae

Footnote 17.

Equation

* I am grateful to Jon Dolling for obtaining the following rather complicated ~~the~~ solution for equations 1 to 7 and 3.

$$\Omega_A = \frac{B + R(1-B)}{AB + R(1-AB)}$$

where R is the unique root in the range $0 < R < 1$ of the following quadratic equation

$$R^2 \left\{ (1-A)(1-B) - (1-AB)^2 \Omega \right\}$$
$$+ R \left\{ A + B - 2AB - 2AB(1-AB)\Omega \right\}$$
$$+ AB(1 - ABR\Omega) = 0 .$$

We leave the derivation of this result as a challenge to the interested reader.

It is conjectured that (16) > (17) ~~are the~~ is the only simple solution satisfying the conditions (1) ~~to~~ (6)

zero hard

The only other property we might like to impose is

(7) $\Omega_A = \Omega_B = \sqrt{\Omega}$ for $A = B$.

Our solution (16) > (17) does not satisfy condition (7) ~~and it may even be the case that conditions 1 to 7 are incompatible~~. There ~~is certainly~~ appears to be no simple solution that satisfies all 7 conditions (1) to (7). * 12 Indeed

Hence we are only interested in a model which will bring out the general features of different methodologies. We shall adopt our solutions (16) and (17) as

well hatched

Another ~~feature~~ feature we might want to build into our model is a memory[13] (Footnote 13: I am grateful to Noella Kenton for suggesting this). That is to say after a number m of unsuccessful predictions we do not choose to leave the probability of the hard core entirely unchanged but "depress" it at the mth repetition by a "memory factor" $F(m)$ which begins to cut Ω at some critical value m_0.

1 line

p 13

We can now put all these features together to obtain a model for the following very general problem

new para

Suppose a research programme makes n_m successful predictions following the $(m-1)^{th}$ false prediction by suitable articulation of its positive heuristic

new para

We denote by $P_c(n_1, n_2 \ldots n_m)$ the posterior probability for the hard core c after this sequence of successes and failures

Then we can write the following recursion for P_c

$$P_c(n_1, n_2, \ldots n_m) = P_c'(n_1 \ldots n_{m-1}) \times \Theta\left(\sqrt{x}, P_c'(n_1 \ldots n_{m-1}), \Omega^{(n_m)}\right)$$

where $P_c'(n_1 \ldots n_{m-1}) = F(m-1) \times P_c(n_1 \ldots n_{m-1})$

and $\Omega^{(n_m)} = \Phi\left(\varepsilon, n_m, \sqrt{x}\, P_c'(n_1 \ldots n_{m-1})\right)$

$\Omega^{(n_1)} = \Phi(\varepsilon, n_1, x)$

with $P_c(n_1) = \sqrt{x}\, \Theta\left(\sqrt{x}, \sqrt{x}, \Omega^{(n_1)}\right)$

and $\Theta(A, B, c) = \dfrac{c(1-AB)}{1-B + B(1-A)c}$

$\Phi(A, B, c) = \cancel{1-A^B} \dfrac{1}{c(1-A^B) + A^B}$

$F(m) = \dfrac{1}{1 + \exp(m - m_0)}$

In these formulae as before $x = p(C \& B_i \mid b)$ and

$$\varepsilon = p(e \mid \sim(C \& B_i)).$$

We have assumed that initially equal probabilities $\sqrt{x}$ are attached to C and B_1 and that each successive articulation of the protective belt starts again with probability $\sqrt{x}$ ~~although what happens to P_c after refutation will depend on our methodology~~.

new para We have chosen a simple form for $F(m)$ with suitable cut-off properties.

new para ~~Of course if the number of alternative hypotheses in the protective belt that can be invoked to deal with anomalies is a finite number N say, then taking the pure Lakatos model without a memory factor, P_c will drop immediately to zero after $N+1$ anomalies have appeared since no further articulation of the protective belt is possible.~~

new para Taking the parameter values $N=6$, $m_0 = 3$, $\sqrt{x} = 0.01$, $\varepsilon = .1$ and assuming for purposes of illustration the case where successful and unsuccessful predictions alternate, the account derived from our model is shown in Fig.3 where curves for pure Lakatos (no memory factor) and pure Popper (refutation takes us back to square one, i.e. P_c falls to $\sqrt{x}$ after each refutation) are also given for comparison.

new para Certainly the memory model looks ok

most realistic although we have chosen a very simple type of money — a more sophisticated version would take due account of small or vanishing values of N_m ~~for~~ as m increases, corresponding to the informal notion of degeneracy in the research programme.

~~At all events we have succeeded in our~~ ~~[illegible] of giving a simple Bayesian model for the life and death of a research programme~~

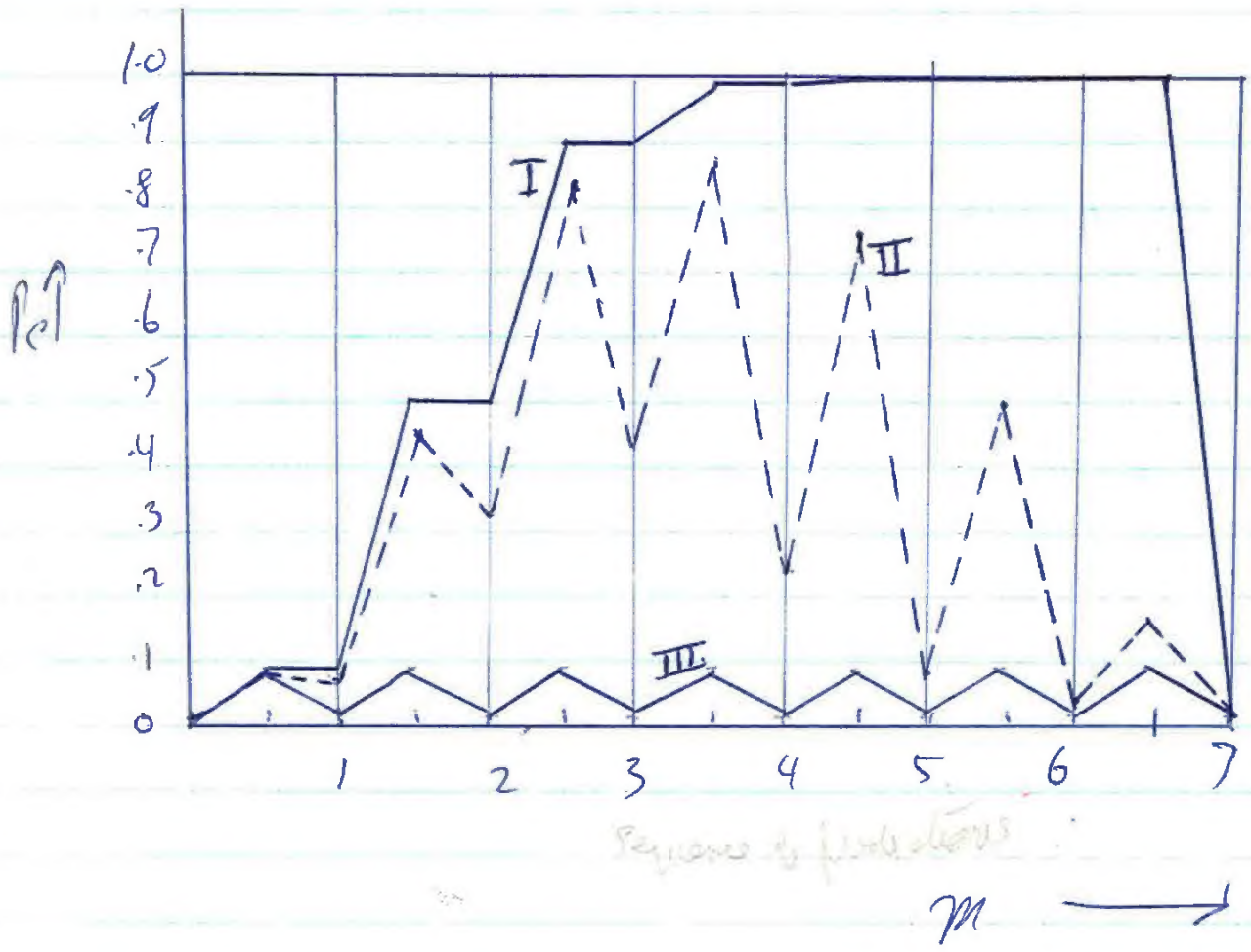


Fig. 3.

$$\left.\begin{aligned} N &= 6 \\ m_0 &= 3 \\ \sqrt{x} &= \cdot 0.1 \\ \varepsilon &= \cdot 1 \end{aligned}\right\}$$

I pure Lakatos : refutation leaves P_c unchanged

II Lakatos plus memory ; refutation cuts P_c by a memory factor $F(m)$.

III Pure Popper : refutation takes P_c back to $\sqrt{x}$.

new paid

Let us now briefly compare our new model with the freedom one presented in Section 5.

In the first place the new model assumes that B_{and} C are independent after refutation – this is explicitly denied on the purely Bayesian model.

Moreover the first model suppose more information about more complicated conditional probabilities such as $P(r_e/B \& r_c)$ etc whereas the second model assumes we know only the value of $P(r_e/r_B \vee r_C)$ as our input information.

Naturally our methodological decisions ~~about how to~~ in the second model could be re-expressed as constraints on the ~~additional~~ extra conditional probability assignments assumed in the first model – (~~this is how Jon Darling arrived at his solution of our constraints to on the for the second model~~ in fact)

But the status of these assignments is now quite different. For the first model we denied the effect of refutation from these assignments. In the second model we effectively constrain these assignments from by the decisions we choose to ~~take~~ make about the effects of refutation.

Finally we note that in the second model we can explain the progress of the rational programme ~~we~~ even allowing P_B and P_C initially to have the same low value, whereas the success of the first

model depends critically on taking these as zero-goal.
At all events we believe that we have successfully distinguished the possible roles of Bayesian or semi-Bayesian models in accounting for the life and death of a research programme.

References Lakatos I [1970]:

Redhead, M.L.G. [1974]: 'On Neyman's Paradox and the theory of statistical Tests' Brit. J. Phil. Sci, 25 pp. 265-271.

Salmon, W.C. [1967]: The Foundations of Scientific Inference.

Schaffner, K.F. [1974] 'Einstein versus Lorentz: Research Programmes and the Logic of Comparative Theory Evaluation'. Brit. J. Phil Sci 25 pp. 45-78

Zahar, E.G. [1973]: 'Why did Einstein's Programme supersede Lorentz's?' Brit. J. Phil. Sci, 24 pp. 95-123 and 223-62.

Lautrup B.E., Peterman, A. and de Rafael, E. [1972]: 'Comparison between theory, experiment in quantum Electrodynamics' Physics Reports, 3 C (1972), 195-259.

Rich, A. and Wesley J.C. [1972] 'The current status of the lepton g factors' ~~Rev Mod. Phys.~~ (Reviews of Modern Physics) 44, (~~1972~~) pp. 250-283.

Lakatos I [1970]: 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (eds): Criticism and the Growth of Knowledge, pp 91-195.

[Please Marie-Louise write Brit. J. Phil Sci in full (Thus:- The British Journal for the Philosophy of Science - in references for Redhead [1974], Schaffner [1974] and Zahar [1973]]